

Optimal control of forward-backward stochastic Volterra equations

Nacira Agram^{1,2}, Bernt Øksendal^{1,2} and Samia Yakhlef³

10 March 2017

Abstract: We study the problem of optimal control of a coupled system of forward-backward stochastic Volterra equations. We use Hida-Malliavin calculus to prove a sufficient and a necessary maximum principle for the optimal control of such systems. Existence and uniqueness of backward stochastic Volterra integral equations are proved. As an application of our methods, we solve a recursive utility optimisation problem in a financial model with memory.

Keywords: Forward-backward stochastic Volterra equation, optimal control, partial information, Hida-Malliavin calculus, maximum principles, optimal recursive utility consumption.

MSC 2010: 60H20, 60G57, 60J70, 60J75, 93E20, 91B28, 91B70, 91B42.

1 Introduction

The purpose of this paper is to establish solution techniques for optimal control of coupled systems of stochastic Volterra equations. Stochastic Volterra equations appear in models for dynamic systems with noise and memory. As a motivating example, consider the following Volterra equation, modelling a stochastic cash flow $X(t) = X^c(t)$ subject to a consumption rate $c(t)$ at time t :

$$X(t) = \xi(t) + \int_0^t (\alpha(t, s) - c(s)) X(s) ds + \int_0^t \beta(t, s) X(s) dB(s) + \int_0^t \int_{\mathbb{R}} \pi(t, s, e) X(s) \tilde{N}(ds, de), t \in [0, T], \quad (1.1)$$

¹Department of Mathematics, University of Oslo, P.O. Box 1053 Blindern, N-0316 Oslo, Norway.
Email: naciraa@math.uio.no, oksendal@math.uio.no

²This research was carried out with support of the Norwegian Research Council, within the research project Challenges in Stochastic Control, Information and Applications (STOCONINF), project number 250768/F20.

³Department of Mathematics, University of Biskra, Algeria.
Email: samiayakhlef@yahoo.fr

where $\xi : [0, T] \rightarrow \mathbb{R}$ and $\alpha, \beta : [0, T]^2 \rightarrow \mathbb{R}$ and $\pi : [0, T]^2 \times \mathbb{R}_0 \rightarrow \mathbb{R}$ are deterministic functions with α, β and π bounded.

Here $B(t) = B(t, \omega)$ and $N(dt, de) = N(dt, de, \omega)$ are a Brownian motion and an independent Poisson random measure, respectively, on a complete probability space (Ω, \mathcal{F}, P) . The compensated Poisson random measure \tilde{N} is defined by $\tilde{N}(dt, de) = N(dt, de) - \nu(de)dt$, where ν is the Lévy measure of N . We denote by $\mathbb{F} = \{\mathcal{F}_t\}_{t \geq 0}$ the right-continuous complete filtration generated by B and N and we let

$$\mathbb{G} := \{\mathcal{G}_t\}_{t \geq 0}$$

be a given right-continuous complete subfiltration of \mathbb{F} , in the sense that

$$\mathcal{G}_t \subseteq \mathcal{F}_t \text{ for all } t \in [0, T].$$

The sigma-algebra \mathcal{G}_t represents the information available to the consumer at time t . Let $\mathcal{P}(\mathbb{F})$ be the σ -algebra of \mathbb{F} -predictable subsets of $\Omega \times \mathbb{R}_+$, i.e., the σ -algebra generated by the left continuous \mathbb{F} -adapted processes. The *forward stochastic Volterra integral equation* (FSVIE) (1.1) can be written in differential form as

$$\begin{aligned} dX(t) = & \xi'(t)dt + (\alpha(t, t) - c(t))X(t)dt + \left(\int_0^t \frac{\partial \alpha}{\partial t}(t, s)X(s)ds \right) dt \\ & + \beta(t, t)X(t)dB(t) + \left(\int_0^t \frac{\partial \beta}{\partial t}(t, s)X(s)dB(s) \right) dt \\ & + \int_{\mathbb{R}} \pi(t, t, e)X(t)\tilde{N}(dt, de) \\ & + \left(\int_{\mathbb{R}} \int_0^t \frac{\partial \pi}{\partial t}(t, s, e)X(s)\tilde{N}(ds, de) \right) dt, t \in [0, T]. \end{aligned} \quad (1.2)$$

From (1.2) we see that the dynamics of $X(t)$ contains history or memory terms represented by the ds -integrals.

Following a suggestion of Duffie and Epstein [5] we now model the total utility of the consumption rate $c(t)$ by a *recursive utility* process $Y(t) = Y^c(t)$ defined by the equation

$$Y(t) = \mathbb{E} \left[- \int_t^T \{ \gamma(s)Y(s) + \ln(c(s)X(s)) \} ds \mid \mathcal{F}_t \right]; \quad t \in [0, T] \quad (1.3)$$

By the martingale representation theorem we see that there exist processes $Z(t), K(t, e)$ such that the triple (Y, Z, K) solves the backward stochastic differential equation (BSDE)

$$\begin{cases} dY(t) = - [\gamma(t)Y(t) + \ln(c(t)X(t))] dt + Z(t)dB(t) \\ \quad + \int_{\mathbb{R}} K(t, e)\tilde{N}(dt, de); \quad t \in [0, T] \\ Y(T) = 0 \end{cases} \quad (1.4)$$

We now consider the *optimal recursive utility problem* to maximise the total recursive utility of the consumption. In other words, we want to find an optimal consumption rate $c^* \in \mathcal{U}_{\mathbb{G}}$ such that

$$\sup_{c \in \mathcal{U}_{\mathbb{G}}} Y^c(0) = Y^{c^*}(0), \quad (1.5)$$

where $\mathcal{U}_{\mathbb{G}}$ is a given set of admissible \mathbb{G} -adapted consumption processes.

This is a problem of optimal control of a coupled system consisting of the forward stochastic Volterra equation (1.1) and the BSDE (1.4). In the following sections we will present solution methods for general optimal control for systems of forward-backward stochastic Volterra equations. Then in the last section we will apply the methods to solve the optimal recursive utility consumption problem above.

There has been a lot of research activity recently within stochastic Volterra integral equations (SVIEs) recently, both of forward and backward type. See e.g. [2], [7], [10], [11], [13], [12], [14], [15], [16] and [17]. Perhaps the paper closest to our paper is [12]. However, that paper has a different approach than our paper, does not have a sufficient maximum principle and does not deal with jumps and partial information, as we do.

2 Stochastic maximum principle for FBSVE

This section is an extension to forward-backward systems of the results obtained in [2]. We consider a system governed by a coupled system of controlled forward-backward stochastic Volterra equations (FBSVE) of the form:

$$X(t) = \xi(t) + \int_0^t b(t, s, X(s), u(s))ds + \int_0^t \sigma(t, s, X(s), u(s))dB(s) + \int_0^t \int_{\mathbb{R}} \theta(t, s, X(s), u(s), e) \tilde{N}(ds, de), t \in [0, T], \quad (2.1)$$

$$Y(t) = \eta(X(T)) + \int_t^T g(t, s, X(s), Y(s), Z(t, s), K(t, s, \cdot), u(s))ds - \int_t^T Z(t, s)dB(s) - \int_t^T \int_{\mathbb{R}} K(t, s, e) \tilde{N}(ds, de), t \in [0, T]. \quad (2.2)$$

The quadruple (X, Y, Z, K) is said to be a solution of (2.1)-(2.2) if it satisfies both equations. To the best of our knowledge, results about existence and uniqueness of solutions for such general systems are not known. Conditions under which there exists a unique solution (Y, Z, K) of (2.2) are studied in section 3.

In the above the functions ξ, η are assumed to be deterministic and C^1 , while the functions

$$\begin{aligned} b(t, s, x, u) &: [0, T]^2 \times \mathbb{R} \times \mathbb{U} \times \Omega \rightarrow \mathbb{R}, \\ \sigma(t, s, x, u) &: [0, T]^2 \times \mathbb{R} \times \mathbb{U} \times \Omega \rightarrow \mathbb{R}, \\ g(t, s, x, y, z, k(\cdot), u) &: [0, T]^2 \times \mathbb{R}^3 \times L^2(\nu) \times \mathbb{U} \times \Omega \rightarrow \mathbb{R}, \\ \theta(t, s, x, u, e) &: [0, T]^2 \times \mathbb{R} \times \mathbb{U} \times \mathbb{R}_0 \times \Omega \rightarrow \mathbb{R}, \end{aligned}$$

are assumed to be continuously differentiable with respect to their first variables, and for all t, x, y, z, k, u, e the processes $s \mapsto b(t, s, x, u), s \mapsto \sigma(t, s, x, u), s \mapsto g(t, s, x, y, z, k(\cdot), u), s \mapsto \theta(t, s, x, u, e)$ are \mathcal{F}_s -measurable for all $s \leq t$. We assume that $t \mapsto Z(t, s)$ and $t \mapsto K(t, s, \cdot)$ are C^1 for all s, e, ω and that

$$\mathbb{E} \left[\int_0^T \int_0^T \left(\frac{\partial Z}{\partial t}(t, s) \right)^2 dsdt + \int_0^T \int_0^T \int_{\mathbb{R}} \left(\frac{\partial K}{\partial t}(t, s, e) \right)^2 \nu(de) dsdt \right] < \infty. \quad (2.3)$$

It is known that (2.3) holds for some linear systems. See [6].

Let \mathbb{U} be a given open convex subset of \mathbb{R} and let $\mathcal{U} = \mathcal{U}_{\mathbb{G}}$ be a given family of *admissible* controls, required to be \mathbb{G} -predictable, where, as before, $\mathbb{G} = \{\mathcal{G}_t\}_{t \geq 0}$ is a given subfiltration of $\mathbb{F} = \{\mathcal{F}_t\}_{t \geq 0}$, in the sense that $\mathcal{G}_t \subseteq \mathcal{F}_t$ for all t . We associate to the system (2.1) – (2.2) the following *performance functional*:

$$J(u) = \mathbb{E} \left[\int_0^T f(s, X(s), Y(s), u(s)) ds + \varphi(X(T)) + \psi(Y(0)) \right], \quad (2.4)$$

for given functions

$$\begin{aligned} f &: [0, T] \times \mathbb{R}^2 \times \mathbb{U} \times \Omega \rightarrow \mathbb{R}, \\ \varphi &: \mathbb{R} \rightarrow \mathbb{R}, \\ \psi &: \mathbb{R} \rightarrow \mathbb{R}. \end{aligned}$$

The functions φ, ψ are assumed to be C^1 , while $f(s, x, y, u)$ is assumed to be \mathbb{F} -adapted with respect to s and C^1 with respect to x, y, u for each s . We remark here that our performance functional is not of Volterra type. Our optimisation control problem is to find $u^* \in \mathcal{U}_{\mathbb{G}}$ such that

$$\sup_{u \in \mathcal{U}} J(u) = J(u^*). \quad (2.5)$$

Let \mathcal{L} be the set of all \mathbb{F} -adapted stochastic processes, and let \mathcal{R} denote the set of all functions $k : \mathbb{R} \rightarrow \mathbb{R}$. Define the *Hamiltonian functional*:

$$\begin{aligned} \mathcal{H}(t, x, y, z, k(\cdot), v, p, p(\cdot), q, \lambda, \lambda(\cdot), r(\cdot)) &:= H_0(t, x, y, z, k(\cdot), v, p, q, \lambda, r(\cdot)) \\ &+ H_1(t, x, y, z, k(\cdot), v, p(\cdot), \lambda(\cdot)), \end{aligned} \quad (2.6)$$

where

$$H_0 : [0, T] \times \mathbb{R}^3 \times \mathcal{R} \times \mathbb{U} \times \mathbb{R}^3 \times \mathcal{R} \rightarrow \mathbb{R}$$

is defined by

$$\begin{aligned} &H_0(t, x, y, z, k(\cdot), v, p, q, \lambda, r(\cdot)) \\ &:= f(t, x, y, v) + b(t, t, x, v)p + \sigma(t, t, x, v)q \\ &+ \int_{\mathbb{R}} \theta(t, t, x, v)r(t, e)\nu(de) + g(t, t, x, y, z, k(\cdot), v)\lambda \end{aligned} \quad (2.7)$$

and

$$H_1 : [0, T] \times \mathbb{R}^3 \times \mathcal{R} \times \mathbb{U} \times \mathcal{L} \times \mathcal{L} \rightarrow \mathbb{R}$$

is defined by

$$\begin{aligned}
H_1(t, x, y, z, k(\cdot), v, p(\cdot), \lambda(\cdot)) &:= \int_t^T \frac{\partial b}{\partial s}(s, t, x, v) p(s) ds + \int_t^T \frac{\partial \sigma}{\partial s}(s, t, x, v) \mathbb{E}[D_t p(s) | \mathcal{F}_t] ds \\
&+ \int_t^T \int_{\mathbb{R}} \frac{\partial \theta}{\partial s}(s, t, x, v) \mathbb{E}[D_{t,e} p(s) | \mathcal{F}_t] \nu(de) ds + \int_0^t \frac{\partial g}{\partial s}(s, t, x, y, z, k(\cdot), v) \lambda(s) ds \\
&+ \int_0^t \frac{\partial g}{\partial z}(s, t, x, y, z, k(\cdot), v) \frac{\partial Z}{\partial s}(s, t) \lambda(s) ds \\
&+ \int_0^t \left\langle \nabla_k g(s, t, x, y, z, k(\cdot), v), \frac{\partial K}{\partial s}(s, t, \cdot) \right\rangle \lambda(s) ds.
\end{aligned} \tag{2.8}$$

Here, and in the following, D_t and $D_{t,e}$ denote the (*generalised*) *Hida-Malliavin derivative* at t and at (t, e) with respect to B and \tilde{N} , respectively, and ∇_k denotes the *Fréchet derivative* with respect to k . We refer to the Appendix for more details.

The associated *forward-backward system* for the adjoint processes $\lambda(t)$, $(p(t), q(t), r(t, \cdot))$ is

$$\begin{cases} d\lambda(t) := \frac{\partial \mathcal{H}}{\partial y}(t) dt + \frac{\partial \mathcal{H}}{\partial z}(t) dB(t) + \int_{\mathbb{R}} \frac{d\nabla_k \mathcal{H}}{d\nu}(t) \tilde{N}(dt, de), & 0 \leq t \leq T, \\ \lambda(0) := \psi'(Y(0)), \end{cases} \tag{2.9}$$

and

$$\begin{cases} dp(t) := -\frac{\partial \mathcal{H}}{\partial x}(t) dt + q(t) dB(t) + \int_{\mathbb{R}} r(t, e) \tilde{N}(dt, de), & 0 \leq t \leq T, \\ p(T) := \varphi'(X(T)) + \lambda(T) \eta'(X(T)), \end{cases} \tag{2.10}$$

where we have used the simplified notation

$$\frac{\partial \mathcal{H}}{\partial x}(t) = \left[\frac{\partial \mathcal{H}}{\partial x}(t, x, Y(t), Z(t, \cdot), K(t, \cdot), u(t), p(t), q(t), \lambda(t), r(t, \cdot)) \right]_{x=X(t)}, \tag{2.11}$$

and similarly for $\frac{\partial \mathcal{H}}{\partial y}(t)$, $\frac{\partial \mathcal{H}}{\partial z}(t)$...

As in [OS1] we assume that H is Fréchet differentiable (C^1) in the variables x, y, z, k, u and that the Fréchet derivative $\nabla_k H$ of H with respect to $k \in \mathcal{R}$ as a random measure is absolutely continuous with respect to ν , with Radon-Nikodym derivative $\frac{d\nabla_k H}{d\nu}$. Thus, if $\langle \nabla_k H, h \rangle$ denotes the action of the linear operator $\nabla_k H$ on the function $h \in \mathcal{R}$ we have

$$\langle \nabla_k H, h \rangle = \int_{\mathbb{R}} h(\zeta) d\nabla_k H(\zeta) = \int_{\mathbb{R}} h(\zeta) \frac{d\nabla_k H(\zeta)}{d\nu(\zeta)} d\nu(\zeta). \tag{2.12}$$

The question of existence and uniqueness of the forward-backward system above will not be studied here. It is a subject of future research. See, however our partial result in Section 3.

2.1 A sufficient maximum principle

In this subsection, we prove that under some conditions such as the concavity, a given control \hat{u} which satisfies a maximum condition of the Hamiltonian, is an optimal control for the problem (2.5).

From (2.1) – (2.2) we can get the differential forms:

$$\begin{aligned} dX(t) = & \xi'(t)dt + b(t, t, X(t), u(t))dt + \left(\int_0^t \frac{\partial b}{\partial t}(t, s, X(s), u(s))ds \right) dt \\ & + \sigma(t, t, X(t), u(t))dB(t) + \left(\int_0^t \frac{\partial \sigma}{\partial t}(t, s, X(s), u(s))dB(s) \right) dt \\ & + \int_{\mathbb{R}} \theta(t, t, X(s), u(s), e)\tilde{N}(ds, de) + \left(\int_0^t \int_{\mathbb{R}} \frac{\partial \theta}{\partial t}(t, s, X(s), u(s), e)\tilde{N}(ds, de) \right) dt, \end{aligned} \quad (2.13)$$

and

$$\begin{aligned} dY(t) = & -g(t, t, X(t), Y(t), Z(t, t), K(t, t, \cdot), u(t))dt \\ & + \left(\int_t^T \frac{\partial g}{\partial t}(t, s, X(s), Y(s), Z(t, s), K(t, s, \cdot), u(s))ds \right) dt \\ & + \int_t^T \frac{\partial g}{\partial z}(t, s, X(s), Y(s), Z(t, s), K(t, s, \cdot), u(s)) \frac{\partial Z}{\partial t}(t, s) dt \\ & + \int_t^T \left\langle \nabla_k g(t, s, X(s), Y(s), Z(t, s), K(t, s, \cdot), u(s)), \frac{\partial K}{\partial t}(t, s, \cdot) \right\rangle dt \\ & + Z(t, t)dB(t) + \int_{\mathbb{R}} K(t, t, e)\tilde{N}(dt, de) \\ & - \left(\int_t^T \frac{\partial Z}{\partial t}(t, s)dB(s) \right) dt - \left(\int_t^T \int_{\mathbb{R}} \frac{\partial K}{\partial t}(t, s, e)\tilde{N}(ds, de) \right) dt, \\ Y(T) = & \eta(X(T)). \end{aligned} \quad (2.14)$$

We now state and prove a sufficient maximum principle:

Theorem 2.1. *Let $\hat{u} \in \mathcal{U}_{\mathbb{G}}$, with corresponding solutions $\hat{X}(t)$, $(\hat{Y}(t), \hat{Z}(t, s), \hat{K}(t, s, \cdot), \hat{\lambda}(t), (\hat{p}(t), \hat{q}(t), \hat{r}(t, \cdot)))$ of equations (2.13), (2.14), (2.9) and (2.10), respectively. Assume the following:*

- (Concavity conditions) *The functions*

$$x \mapsto \eta(x), \quad x \mapsto \varphi(x), \quad x \mapsto \psi(x)$$

and

$$x, y, z, k(\cdot), u \mapsto \mathcal{H}(t, x, y, z, k(\cdot), u, p, q, \lambda, r),$$

are concave for all t, p, q, λ, r .

- (The maximum condition)

$$\begin{aligned} & \sup_{v \in \mathcal{U}} \mathbb{E} \left[\mathcal{H}(t, \hat{X}(t), \hat{Y}(t), \hat{Z}(t), \hat{k}(t, \cdot), v, \hat{\lambda}(t), \hat{p}(t), \hat{q}(t), \hat{r}(t, \cdot)) \mid \mathcal{G}_t \right] \\ &= \mathbb{E} \left[\mathcal{H}(t, \hat{X}(t), \hat{Y}(t), \hat{Z}(t), \hat{k}(t, \cdot), \hat{u}(t), \hat{\lambda}(t), \hat{p}(t), \hat{q}(t), \hat{r}(t, \cdot)) \mid \mathcal{G}_t \right], \quad \forall t \geq 0. \end{aligned} \quad (2.15)$$

Then, \hat{u} is an optimal \mathbb{G} -adapted control.

Proof. By considering a suitable increasing family of stopping times converging to T , we may assume that all the local martingales appearing in the proof below are martingales. In particular, the expectations of the dB - and $\tilde{N}(dt, de)$ -integrals are all 0.

Choose an arbitrary $u \in \mathcal{U}_{\mathbb{G}}$ and consider

$$J(u) - J(\hat{u}) = I_1 + I_2 + I_3,$$

where

$$\begin{aligned} I_1 &= \mathbb{E} \left[\int_0^T \left\{ f(t) - \hat{f}(t) \right\} dt \right], \quad I_2 = \mathbb{E} \left[\varphi(X(T)) - \varphi(\hat{X}(T)) \right], \\ I_3 &= \mathbb{E} \left[\psi(Y(0)) - \psi(\hat{Y}(0)) \right], \end{aligned} \quad (2.16)$$

where $f(t) = f(t, X(t), Y(t), u(t))$, $\hat{f}(t) = f(t, \hat{X}(t), \hat{Y}(t), \hat{u}(t))$. Using a simplified notation

$$\begin{aligned} b(t, t) &= b(t, t, X(t), u(t)), \hat{b}(t, t) = b(t, t, \hat{X}(t), \hat{u}(t)), b(t, s) = b(t, s, X(s), u(s)) \\ \theta(t, t, e) &= \theta(t, t, X(s), u(s), e), \theta(t, s, e) = \theta(t, s, X(s), u(s), e) \text{ etc., we get} \end{aligned}$$

$$\begin{aligned} I_1 &= \mathbb{E} \left[\int_0^T \left\{ H_0(t) - \hat{H}_0(t) - \hat{p}(t) \left(b(t, t) - \hat{b}(t, t) \right) - \hat{q}(t) \left(\sigma(t, t) - \hat{\sigma}(t, t) \right) \right. \right. \\ &\quad \left. \left. - \hat{\lambda}(t) \left(g(t, t) - \hat{g}(t, t) \right) - \int_{\mathbb{R}} \hat{r}(t, e) \left(\theta(t, t, e) - \hat{\theta}(t, t, e) \right) \nu(de) \right\} dt \right]. \end{aligned} \quad (2.17)$$

Using concavity and the Itô formula, we obtain

$$\begin{aligned}
I_2 &\leq \mathbb{E} \left[\varphi'(\hat{X}(T)) \left(X(T) - \hat{X}(T) \right) \right] \\
&= \mathbb{E} \left[\hat{p}(T) \left(X(T) - \hat{X}(T) \right) \right] - \mathbb{E} \left[\hat{\lambda}(T) \eta'(\hat{X}(T)) \left(X(T) - \hat{X}(T) \right) \right] \\
&= \mathbb{E} \left[\int_0^T \hat{p}(t) \left(dX(t) - d\hat{X}(t) \right) + \int_0^T \left(X(t) - \hat{X}(t) \right) d\hat{p}(t) \right. \\
&\quad \left. + \int_0^T \hat{q}(t) (\sigma(t, t) - \hat{\sigma}(t, t)) dt + \int_0^T \int_{\mathbb{R}} \hat{r}(t, e) (\theta(t, t, e) - \hat{\theta}(t, t, e)) \nu(de) dt \right] \\
&\quad - \mathbb{E} \left[\hat{\lambda}(T) \eta'(\hat{X}(T)) \left(X(T) - \hat{X}(T) \right) \right] \\
&= \mathbb{E} \left[\int_0^T \left\{ \hat{p}(t) \left(b(t, t) - \hat{b}(t, t) + \int_0^t \left(\frac{\partial b}{\partial t}(t, s) - \frac{\partial \hat{b}}{\partial t}(t, s) \right) ds \right. \right. \right. \\
&\quad \left. \left. + \int_0^t \left(\frac{\partial \sigma}{\partial t}(t, s) - \frac{\partial \hat{\sigma}}{\partial t}(t, s) \right) dB(s) \right. \right. \\
&\quad \left. \left. + \int_0^t \int_{\mathbb{R}} \left(\frac{\partial \theta}{\partial t}(t, s) - \frac{\partial \hat{\theta}}{\partial t}(t, s) \right) \tilde{N}(ds, de) \right) - \frac{\partial \hat{\mathcal{H}}}{\partial x}(t) (X(t) - \hat{X}(t)) \right. \\
&\quad \left. + \hat{q}(t) [\sigma(t, t) - \hat{\sigma}(t, t)] \right\} dt \right] + \int_0^T \int_{\mathbb{R}} \hat{r}(t, e) (\theta(t, t, e) - \hat{\theta}(t, t, e)) \nu(de) dt \\
&\quad - \mathbb{E} [\hat{\lambda}(T) \eta'(\hat{X}(T)) (X(T) - \hat{X}(T))]. \tag{2.18}
\end{aligned}$$

By the Fubini theorem, we get

$$\begin{aligned}
\int_0^T \left(\int_0^t \frac{\partial b}{\partial t}(t, s) ds \right) \hat{p}(t) dt &= \int_0^T \left(\int_s^T \frac{\partial b}{\partial t}(t, s) \hat{p}(t) dt \right) ds \\
&= \int_0^T \left(\int_t^T \frac{\partial b}{\partial s}(s, t) \hat{p}(s) ds \right) dt, \tag{2.19}
\end{aligned}$$

and similarly, by the generalised duality theorems for the Malliavin derivatives [2], we have

$$\begin{aligned}
\mathbb{E} \left[\int_0^T \left(\int_0^t \frac{\partial \sigma}{\partial t}(t, s) dB(s) \right) \hat{p}(t) dt \right] &= \int_0^T \mathbb{E} \left[\int_0^t \frac{\partial \sigma}{\partial t}(t, s) dB(s) \hat{p}(t) \right] dt \\
&= \int_0^T \mathbb{E} \left[\int_0^t \frac{\partial \sigma}{\partial t}(t, s) \mathbb{E}[D_s \hat{p}(t) \mid \mathcal{F}_s] ds \right] dt \\
&= \int_0^T \mathbb{E} \left[\int_s^T \frac{\partial \sigma}{\partial t}(t, s) \mathbb{E}[D_s \hat{p}(t) \mid \mathcal{F}_s] dt \right] ds \\
&= \mathbb{E} \left[\int_0^T \int_t^T \frac{\partial \sigma}{\partial s}(s, t) \mathbb{E}[D_t \hat{p}(s) \mid \mathcal{F}_t] ds dt \right] \tag{2.20}
\end{aligned}$$

and

$$\begin{aligned}
& \mathbb{E} \left[\int_0^T \left(\int_0^t \int_{\mathbb{R}} \left(\frac{\partial \theta}{\partial t}(t, s) \right) \tilde{N}(ds, de) p(t) \right) dt \right] \\
&= \int_0^T \mathbb{E} \left[\int_0^t \int_{\mathbb{R}} \left(\frac{\partial \theta}{\partial t}(t, s) \right) \tilde{N}(ds, de) p(t) \right] dt \\
&= \int_0^T \mathbb{E} \left[\int_s^T \int_{\mathbb{R}} \frac{\partial \theta}{\partial t}(t, s) \mathbb{E} [D_{s,e} p(t) | \mathcal{F}_s] \nu(de) dt \right] ds \\
&= \int_0^T \mathbb{E} \left[\int_t^T \int_{\mathbb{R}} \frac{\partial \theta}{\partial s}(s, t) \mathbb{E} [D_{t,e} p(s) | \mathcal{F}_t] \nu(de) ds \right] dt
\end{aligned} \tag{2.21}$$

Substituting (2.20), (2.21) and (2.19) into (2.18), we get

$$\begin{aligned}
I_2 \leq & \mathbb{E} \left[\int_0^T \left\{ \hat{p}(t) \left(b(t, t) - \hat{b}(t, t) \right) + \int_t^T \hat{p}(s) \left(\frac{\partial b}{\partial s}(s, t) - \frac{\partial \hat{b}}{\partial s}(s, t) \right) ds \right. \right. \\
& + \int_t^T \left(\frac{\partial \sigma}{\partial s}(s, t) - \frac{\partial \hat{\sigma}}{\partial s}(s, t) \right) \mathbb{E}[D_t \hat{p}(s) | \mathcal{F}_t] ds \\
& + \int_t^T \int_{\mathbb{R}} \frac{\partial \theta}{\partial s}(s, t, e) \mathbb{E} [D_{t,e} p(s) | \mathcal{F}_t] \nu(de) ds \\
& + \int_0^T \int_{\mathbb{R}} \left(\hat{r}(t, e) \left(\theta(t, t, e) - \hat{\theta}(t, t, e) \right) \nu(de) \right) dt \\
& \left. - \frac{\partial \hat{\mathcal{H}}}{\partial x}(t) \left(X(t) - \hat{X}(t) \right) + \hat{q}(t) \left(\sigma(t, t) - \hat{\sigma}(t, t) \right) \right\} dt \right] \\
& - \mathbb{E} \left[\hat{\lambda}(T) \eta'(\hat{X}(T)) \left(X(T) - \hat{X}(T) \right) \right].
\end{aligned} \tag{2.22}$$

By the concavity of ψ and η , we obtain

$$\begin{aligned}
I_3 &= \mathbb{E} \left[\psi(Y(0)) - \psi(\hat{Y}(0)) \right] \\
&\leq \mathbb{E} \left[\psi'(\hat{Y}(0)) \left(Y(0) - \hat{Y}(0) \right) \right] \\
&= \mathbb{E} \left[\hat{\lambda}(0) \left(Y(0) - \hat{Y}(0) \right) \right]
\end{aligned}$$

$$\begin{aligned}
&= \mathbb{E} \left[\hat{\lambda}(T) \left(Y(T) - \hat{Y}(T) \right) \right] - \mathbb{E} \left[\int_0^T \left(Y(t) - \hat{Y}(t) \right) d\hat{\lambda}(t) \right. \\
&+ \int_0^T \hat{\lambda}(t) \left(dY(t) - d\hat{Y}(t) \right) + \int_0^T \frac{\partial \hat{\mathcal{H}}}{\partial z}(t) \left(Z(t, s) - \hat{Z}(t, s) \right) dt \\
&+ \left. \int_0^T \int_{\mathbb{R}} \frac{d\nabla_k \mathcal{H}}{d\nu}(t) (K(t, s, e) - \hat{K}(t, s, e)) \nu(de) dt \right] \\
&\leq \mathbb{E} \left[\hat{\lambda}(T) \eta'(X(T)) \left(X(T) - \hat{X}(T) \right) \right] \\
&- \mathbb{E} \left[\int_0^T \frac{\partial \hat{\mathcal{H}}}{\partial y}(t) \left(Y(t) - \hat{Y}(t) \right) dt - \int_0^T \hat{\lambda}(t) (g(t, t) - \hat{g}(t, t)) dt \right. \\
&+ \int_0^T \left(\hat{\lambda}(t) \int_t^T \left(\frac{\partial g}{\partial t}(t, s) - \frac{\partial \hat{g}}{\partial t}(t, s) \right) ds \right) dt \\
&+ \int_0^T \hat{\lambda}(t) \left[\int_t^T \left(\frac{\partial g}{\partial z}(t, s) \frac{\partial Z}{\partial t}(t, s) - \frac{\partial \hat{g}}{\partial z}(t, s) \frac{\partial \hat{Z}}{\partial t}(t, s) \right) ds \right] dt \\
&+ \int_0^T \hat{\lambda}(t) \left[\int_t^T \left(\left\langle \nabla_k g(t, s), \frac{\partial K}{\partial t}(t, s, \cdot) \right\rangle - \left\langle \nabla_k \hat{g}(t, s), \frac{\partial \hat{K}}{\partial t}(t, s, \cdot) \right\rangle \right) ds \right] dt \\
&+ \int_0^T \left(\hat{\lambda}(t) \int_t^T \left(\frac{\partial Z}{\partial t}(t, s) - \frac{\partial \hat{Z}}{\partial t}(t, s) \right) dB(s) \right) dt \\
&+ \int_0^T \left(\hat{\lambda}(t) \int_t^T \int_{\mathbb{R}} \left(\frac{\partial K}{\partial t}(t, s, \cdot) - \frac{\partial \hat{K}}{\partial t}(t, s, \cdot) \right) \tilde{N}(ds, de) \right) dt \\
&+ \int_0^T \frac{\partial \hat{\mathcal{H}}}{\partial z}(t) \left(Z(t, s) - \hat{Z}(t, s) \right) dt \\
&+ \left. \int_0^T \int_{\mathbb{R}} \frac{d\nabla_k \hat{\mathcal{H}}}{d\nu}(t) \left(K(t, s, e) - \hat{K}(t, s, e) \right) \nu(de) dt \right]. \tag{2.23}
\end{aligned}$$

By the Fubini Theorem, we get

$$\begin{aligned}
\int_0^T \left(\int_t^T \frac{\partial g}{\partial t}(t, s) ds \right) \hat{\lambda}(t) dt &= \int_0^T \left(\int_0^s \frac{\partial g}{\partial t}(t, s) \hat{\lambda}(t) dt \right) ds \\
&= \int_0^T \left(\int_0^t \frac{\partial g}{\partial s}(s, t) \hat{\lambda}(s) ds \right) dt, \tag{2.24}
\end{aligned}$$

and

$$\int_0^T \hat{\lambda}(t) \left[\int_t^T \frac{\partial g}{\partial z}(t, s) \frac{\partial Z}{\partial t}(t, s) ds \right] dt = \int_0^T \left(\int_0^t \hat{\lambda}(s) \frac{\partial g}{\partial z}(s, t) \frac{\partial Z}{\partial s}(s, t) ds \right) dt, \quad (2.25)$$

$$\begin{aligned} & \int_0^T \hat{\lambda}(t) \left[\int_t^T \left\langle \nabla_k g(t, s), \frac{\partial K}{\partial t}(t, s, \cdot) \right\rangle ds \right] dt \\ &= \int_0^T \left(\int_0^t \hat{\lambda}(s) \left\langle \nabla_k g(s, t), \frac{\partial K}{\partial s}(s, t, \cdot) \right\rangle ds \right) dt. \end{aligned} \quad (2.26)$$

Substituting (2.24)-(2.26) into (5.18), we get

$$\begin{aligned} I_3 &\leq \mathbb{E} \left[\hat{\lambda}(T) \eta'(X(T)) (X(T) - \hat{X}(T)) \right] \\ &- \mathbb{E} \left[\int_0^T \frac{\partial \hat{\mathcal{H}}}{\partial y}(t) (Y(t) - \hat{Y}(t)) dt - \int_0^T \hat{\lambda}(t) (g(t, t) - \hat{g}(t, t)) dt \right. \\ &+ \int_0^T \int_0^t \left(\frac{\partial g}{\partial s}(s, t) - \frac{\partial \hat{g}}{\partial s}(s, t) \right) \hat{\lambda}(s) ds dt \\ &+ \int_0^T \left(\int_0^t \hat{\lambda}(s) \left[\frac{\partial g}{\partial z}(s, t) \frac{\partial Z}{\partial s}(s, t) - \frac{\partial \hat{g}}{\partial z}(s, t) \frac{\partial \hat{Z}}{\partial s}(s, t) \right] ds \right) dt \\ &+ \int_0^T \left(\int_0^t \hat{\lambda}(s) \left[\left\langle \nabla_k g(s, t), \frac{\partial K}{\partial s}(s, t, \cdot) \right\rangle \right. \right. \\ &\left. \left. - \left\langle \nabla_k \hat{g}(s, t), \frac{\partial \hat{K}}{\partial s}(s, t, \cdot) \right\rangle \right] ds \right) dt + \int_0^T \frac{\partial \hat{\mathcal{H}}}{\partial z}(t) (Z(t, s) - \hat{Z}(t, s)) dt \Big]. \end{aligned} \quad (2.27)$$

Adding (2.17), (2.22) and (2.27), and noting that

$$\begin{aligned} H_1(t) - \hat{H}_1(t) &= \int_t^T \left\{ \frac{\partial b}{\partial s}(s, t) - \frac{\partial \hat{b}}{\partial s}(s, t) \right\} \hat{p}(s) ds \\ &+ \int_t^T \left\{ \frac{\partial \sigma}{\partial s}(s, t) - \frac{\partial \hat{\sigma}}{\partial s}(s, t) \right\} \mathbb{E}[D_t \hat{p}(s) | \mathcal{F}_t] ds \\ &+ \int_0^t \left\{ \frac{\partial g}{\partial s}(s, t) - \frac{\partial \hat{g}}{\partial s}(s, t) \right\} \hat{\lambda}(s) ds \\ &+ \int_0^T \left(\int_0^t \hat{\lambda}(s) \left[\frac{\partial g}{\partial z}(s, t) \frac{\partial Z}{\partial s}(s, t) - \frac{\partial \hat{g}}{\partial z}(s, t) \frac{\partial \hat{Z}}{\partial s}(s, t) \right] ds \right) dt \\ &+ \int_0^T \left(\int_0^t \hat{\lambda}(s) \left[\left\langle \nabla_k g(s, t), \frac{\partial K}{\partial s}(s, t, \cdot) \right\rangle \right. \right. \\ &\left. \left. - \left\langle \nabla_k \hat{g}(s, t), \frac{\partial \hat{K}}{\partial s}(s, t, \cdot) \right\rangle \right] ds \right) dt \\ &+ \int_t^T \int_{\mathbb{R}} \left(\frac{\partial \theta}{\partial s}(s, t, e) - \frac{\partial \hat{\theta}}{\partial s}(s, t, e) \right) \mathbb{E}[D_{t,e} p(s) | \mathcal{F}_t] \nu(de) ds, \end{aligned}$$

we get

$$\begin{aligned}
J(u) - J(\hat{u}) &= I_1 + I_2 + I_3 \\
&\leq \mathbb{E} \left[\int_0^T \left\{ \mathcal{H}(t) - \hat{\mathcal{H}}(t) - \frac{\partial \hat{\mathcal{H}}}{\partial x}(t) (X(t) - \hat{X}(t)) \right. \right. \\
&\quad \left. \left. - \frac{\partial \hat{\mathcal{H}}}{\partial y}(t) (Y(t) - \hat{Y}(t)) - \frac{\partial \hat{\mathcal{H}}}{\partial z}(t) (Z(t, s) - \hat{Z}(t, s)) \right. \right. \\
&\quad \left. \left. - \int_{\mathbb{R}} \frac{d\nabla_k \hat{\mathcal{H}}}{d\nu}(t) (K(t, s, e) - \hat{K}(t, s, e)) \nu(de) \right\} dt \right]. \tag{2.28}
\end{aligned}$$

By the concavity of \mathcal{H} and the maximum condition (2.15), the proof is complete. \square

2.2 A necessary maximum principle

The concavity condition used in the previous subsection does not always hold in applications. We prove now if $\hat{u} \in \mathcal{U}_{\mathbb{G}}$ is an optimal control for the problem (2.5), then we have the equivalence between being a critical point of $J(u)$ and a critical point of the conditional Hamiltonian.

We start by defining the derivative processes. For each given $t \in [0, T]$, let $\alpha = \alpha(t)$ be a bounded \mathcal{G}_t -measurable random variable, let $\epsilon \in (0, T - t]$ and define

$$\mu(s) := \gamma 1_{[t, t+\epsilon]}(s), \quad s \in [0, T]. \tag{2.29}$$

Assume that

$$\hat{u} + \epsilon \mu \in \mathcal{U} \tag{2.30}$$

for all such μ , and all nonzero ϵ sufficiently small. Then the derivative processes are defined by, writing u for \hat{u} for simplicity from now on,

$$\begin{aligned}
X'(t) &:= \frac{d}{d\epsilon} X^{u+\epsilon\mu}(t) |_{\epsilon=0}, \\
Y'(t) &:= \frac{d}{d\epsilon} Y^{u+\epsilon\mu}(t) |_{\epsilon=0}, \\
Z'(t, s) &:= \frac{d}{d\epsilon} Z^{u+\epsilon\mu}(t, s) |_{\epsilon=0}, \\
K'(t, s, \cdot) &:= \frac{d}{d\epsilon} K^{u+\epsilon\mu}(t, s, \cdot) |_{\epsilon=0}.
\end{aligned}$$

We see that

$$\begin{aligned}
X'(t) &= \int_0^t \left(\frac{\partial b}{\partial x}(t, s) X'(s) + \frac{\partial b}{\partial u}(t, s) \mu(s) \right) ds \\
&+ \int_0^t \left(\frac{\partial \sigma}{\partial x}(t, s) X'(s) + \frac{\partial \sigma}{\partial u}(t, s) \mu(s) \right) dB(s) \\
&+ \int_0^t \int_{\mathbb{R}} \left(\frac{\partial \theta}{\partial x}(t, s, e) X'(s) + \frac{\partial \theta}{\partial u}(t, s, e) \mu(s) \right) \tilde{N}(ds, de)
\end{aligned} \tag{2.31}$$

and

$$\begin{aligned}
Y'(t) &= \eta'(X(T)) X'(T) + \int_t^T \left(\frac{\partial g}{\partial x}(t, s) X'(s) + \frac{\partial g}{\partial y}(t, s) Y'(s) \right. \\
&\quad \left. + \frac{\partial g}{\partial z}(t, s) Z'(t, s) + \langle \nabla_k g(t, s), K'(t, s, \cdot) \rangle + \frac{\partial g}{\partial u}(t, s) \mu(s) \right) ds \\
&\quad - \int_t^T Z'(t, s) dB(s) - \int_t^T \int_{\mathbb{R}} K'(t, s, e) \tilde{N}(ds, de).
\end{aligned} \tag{2.32}$$

Hence

$$\begin{aligned}
dX'(t) &= \left[\frac{\partial b}{\partial x}(t, t) X'(t) + \frac{\partial b}{\partial u}(t, t) \mu(t) \right. \\
&\quad + \int_0^t \left(\frac{\partial^2 b}{\partial t \partial x}(t, s) X'(s) + \frac{\partial^2 b}{\partial t \partial u}(t, s) \mu(s) \right) ds \\
&\quad + \int_0^t \left(\frac{\partial^2 \sigma}{\partial t \partial x}(t, s) X'(s) + \frac{\partial^2 \sigma}{\partial t \partial u}(t, s) \mu(s) \right) dB(s) \\
&\quad + \int_0^t \int_{\mathbb{R}} \left(\frac{\partial^2 \theta}{\partial t \partial x}(t, s, e) X'(s) + \frac{\partial^2 \theta}{\partial t \partial u}(t, s, e) \mu(s) \right) \tilde{N}(ds, de) \Big] dt \\
&\quad + \left(\frac{\partial \sigma}{\partial x}(t, t) X'(t) + \frac{\partial \sigma}{\partial u}(t, t) \mu(t) \right) dB(t) \\
&\quad + \int_{\mathbb{R}} \left(\frac{\partial \theta}{\partial x}(t, t, e) X'(t) + \frac{\partial \theta}{\partial u}(t, t, e) \mu(t) \right) \tilde{N}(dt, de),
\end{aligned} \tag{2.33}$$

and

$$\begin{aligned}
dY'(t) &= -\nabla(g(t, t))(X'(t), Y'(t), Z'(t, t), K'(t, t, \cdot), \mu(t))^t dt \\
&+ \int_t^T \nabla \left(\frac{\partial g}{\partial t}(t, s) \right) (X'(s), Y'(s), Z'(t, s), K'(t, s, \cdot), \mu(s))^t dt \\
&+ \int_t^T \nabla \left(\frac{\partial g}{\partial z}(t, s) \right) (X'(s), Y'(s), Z'(t, s), K'(t, s, \cdot), \mu(s))^t \left(\frac{\partial Z}{\partial t}(t, s) \right) dt \\
&+ \int_t^T \nabla (\nabla_k g(t, s)) (X'(s), Y'(s), Z'(t, s), K'(t, s, \cdot), \mu(s))^t \left(\frac{\partial K}{\partial t}(t, s, \cdot) \right) dt \\
&+ Z'(t, t) dB(t) + \int_{\mathbb{R}} K'(t, t, e) \tilde{N}(dt, de) \\
&- \left(\int_t^T \frac{\partial Z'}{\partial t}(t, s) dB(s) \right) dt - \left(\int_t^T \int_{\mathbb{R}} \frac{\partial K'}{\partial t}(t, s, e) \tilde{N}(dt, de) \right) dt,
\end{aligned} \tag{2.34}$$

where we have denoted by ∇ the partial derivatives w.r.t. x, y, z and u and the Fréchet derivative w.r.t k such that $\nabla = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}, \nabla_k, \frac{\partial}{\partial u} \right)^t$ with the second Fréchet derivative $\nabla_k^2 := \nabla_k \nabla_k$.

Theorem 2.2 (Necessary maximum principle). *Let $\hat{u} \in \mathcal{U}_{\mathbb{G}}$ with corresponding solutions $\hat{X}(t), (\hat{Y}(t), \hat{Z}(t, s), \hat{K}(t, s, \cdot)), \hat{\lambda}(t), (\hat{p}(t), \hat{q}(t), \hat{r}(t, \cdot))$ of equations (2.13), (2.14), (2.9) and (2.10), respectively. Then, the following are equivalent:*

(i)

$$\frac{d}{d\epsilon} J(\hat{u} + \epsilon\mu) |_{\epsilon=0} = 0,$$

for all bounded μ of the form (2.29).

(ii)

$$\mathbb{E} \left[\frac{\partial \mathcal{H}}{\partial u}(t) \mid \mathcal{G}_t \right]_{u=\hat{u}} = 0 \text{ for all } t \in [0, T].$$

Proof. Consider

$$\frac{d}{d\epsilon} J(\hat{u} + \epsilon\mu) |_{\epsilon=0} = I_1 + I_2 + I_3, \quad (2.35)$$

where

$$\begin{aligned} I_1 &= \mathbb{E} \left[\int_0^T \left\{ \frac{\partial f}{\partial x}(t) X'(t) + \frac{\partial f}{\partial y}(t) Y'(t) + \frac{\partial f}{\partial u}(t) \mu(t) \right\} dt \right], \\ I_2 &= \mathbb{E} [\varphi'(X(T)) X'(T)] \\ &= \mathbb{E} [p(T) X'(T)] - \mathbb{E} [\lambda(T) \eta'(X(T) X'(T))], \\ I_3 &= \mathbb{E} [\psi'(Y(0)) Y'(0)]. \end{aligned} \quad (2.36)$$

By the Itô formula

$$\begin{aligned} &\mathbb{E} [p(T) X'(T)] \\ &= \mathbb{E} \left[\int_0^T p(t) \left(\frac{\partial b}{\partial x}(t, t) X'(t) + \frac{\partial b}{\partial u}(t, t) \mu(t) \right) dt \right. \\ &\quad + \int_0^T p(t) \left\{ \int_0^t \left(\frac{\partial^2 b}{\partial t \partial x}(t, s) X'(s) + \frac{\partial^2 b}{\partial t \partial u}(t, s) \mu(s) \right) ds \right\} dt \\ &\quad + \int_0^T p(t) \left\{ \int_0^t \left(\frac{\partial^2 \sigma}{\partial t \partial x}(t, s) X'(s) + \frac{\partial^2 \sigma}{\partial t \partial u}(t, s) \mu(s) \right) dB(s) \right\} dt \\ &\quad + \int_0^T p(t) \left\{ \int_0^t \int_{\mathbb{R}} \left(\frac{\partial^2 \theta}{\partial t \partial x}(t, s, e) X'(s) + \frac{\partial^2 \theta}{\partial t \partial u}(t, s, e) \mu(s) \right) \tilde{N}(ds, de) \right\} dt \\ &\quad - \int_0^T X'(t) \frac{\partial \mathcal{H}}{\partial x}(t) dt + \int_0^T q(t) \left(\frac{\partial \sigma}{\partial x}(t, t) X'(t) + \frac{\partial \sigma}{\partial u}(t, t) \mu(t) \right) dt \\ &\quad \left. + \int_0^T \left(\int_{\mathbb{R}} \left(\frac{\partial \theta}{\partial x}(t, t, e) X'(t) + \frac{\partial \theta}{\partial u}(t, t, e) \mu(t) \right) r(t, e) \nu(de) \right) dt \right]. \end{aligned}$$

From (2.19) and (2.20), we have

$$\begin{aligned}
& \mathbb{E} [p(T) X'(T)] \\
&= \mathbb{E} \left[\int_0^T p(t) \left(\frac{\partial b}{\partial x}(t, t) X'(t) + \frac{\partial b}{\partial u}(t, t) \mu(t) \right) dt \right. \\
&+ \int_0^T \int_t^T p(s) \left\{ \left(\frac{\partial^2 b}{\partial s \partial x}(s, t) X'(t) + \frac{\partial^2 b}{\partial s \partial u}(s, t) \mu(t) \right) ds \right\} dt \\
&+ \int_0^T \left\{ \left(\frac{\partial^2 \sigma}{\partial s \partial x}(s, t) X'(t) + \frac{\partial^2 \sigma}{\partial s \partial u}(s, t) \mu(t) \right) \int_t^T \mathbb{E} [D_t p(s) \mid \mathcal{F}_t] ds \right\} dt \\
&+ \int_0^T \left\{ \int_t^T \int_{\mathbb{R}} \left(\frac{\partial^2 \theta}{\partial s \partial x}(s, t, e) X'(t) + \frac{\partial^2 \theta}{\partial s \partial u}(s, t, e) \mu(t) \right) \mathbb{E} [D_{t,e} p(s) \mid \mathcal{F}_t] \nu(de) \right\} dt \\
&- \int_0^T \frac{\partial \mathcal{H}}{\partial x}(t) X'(t) dt + \int_0^T \left(\frac{\partial \sigma}{\partial x}(t) X'(t) + \frac{\partial \sigma}{\partial u}(t) \mu(t) \right) q(t) dt \\
&+ \int_0^T \left(\int_{\mathbb{R}} \left(\frac{\partial \theta}{\partial x}(t, e) X'(t) + \frac{\partial \theta}{\partial u}(t, e) \mu(t) \right) r(t, e) \nu(de) \right) dt \Bigg]. \tag{2.37}
\end{aligned}$$

By the Itô formula and (2.33)-(2.34), we get

$$\begin{aligned}
& \mathbb{E} [\psi'(Y(0)) Y'(0)] = \mathbb{E} [\lambda(0) Y'(0)] \\
&= \mathbb{E} [\lambda(T) Y'(T)] + \mathbb{E} \left[\int_0^T \lambda(t) \left\{ \nabla g(t, t) (X'(t), Y'(t), Z'(t, t), K'(t, t, \cdot), \mu(t))^t \right. \right. \\
&- \int_t^T \left\{ \nabla \left(\frac{\partial g}{\partial t}(t, s), \frac{\partial g}{\partial z}(t, s) \frac{\partial Z}{\partial t}(t, s), \nabla_k g(t, s) \frac{\partial K}{\partial t}(t, s, \cdot) \right) \right. \\
&\left. \left. (X'(t), Y'(t), Z'(t, t), K'(t, t, \cdot), \mu(t))^t \right\} ds dt \right. \\
&+ \int_0^T \lambda(t) \left(\int_t^T \frac{\partial Z'}{\partial t}(t, s) dB(s) \right) dt + \int_0^T \lambda(t) \left(\int_t^T \int_{\mathbb{R}} \frac{\partial K'}{\partial t}(t, s, e) \tilde{N}(dt, de) \right) dt \\
&- \int_0^T \frac{\partial \mathcal{H}}{\partial y}(t) Y'(t) dt - \int_0^T \frac{\partial \mathcal{H}}{\partial z}(t) Z'(t, s) dt \\
&\left. - \int_0^T \int_{\mathbb{R}} \frac{d \nabla_k \mathcal{H}}{d \nu}(t) K'(t, s, e) \nu(de) dt \right].
\end{aligned}$$

From (2.24)-(2.26) and the Fubini theorem, we have

$$\begin{aligned}
& \mathbb{E} [\psi' (Y (0)) Y' (0)] = \mathbb{E} [\lambda (T) Y' (T)] \\
& + \mathbb{E} \left[\int_0^T \lambda (t) \left\{ \nabla g(t, t) (X' (t), Y' (t), Z' (t, t), K' (t, t, \cdot), \mu (t))^t \right\} \right. \\
& + \int_0^T \int_0^t \lambda (s) \left\{ \nabla \left(\frac{\partial g}{\partial t}(s, t), \frac{\partial g}{\partial z}(s, t) \frac{\partial Z}{\partial t}(s, t), \nabla_k g(s, t) \frac{\partial K}{\partial t}(s, t, \cdot) \right) \right. \\
& \left. \left. (X' (s), Y' (s), Z' (t, s), K' (t, s, \cdot), \mu (s))^t \right\} ds dt \right. \\
& - \int_0^T \frac{\partial \mathcal{H}}{\partial y}(t) Y' (t) dt - \int_0^T \frac{\partial \mathcal{H}}{\partial z}(t) Z' (t, s) dt \\
& \left. - \int_0^T \int_{\mathbb{R}} \frac{d \nabla_k \mathcal{H}}{\partial \nu}(t) K' (t, s, e) \nu (de) dt \right]. \tag{2.38}
\end{aligned}$$

Using that

$$\begin{aligned}
\frac{\partial \mathcal{H}}{\partial x}(t) &= \frac{\partial f}{\partial x}(t) + \frac{\partial b}{\partial x}(t, t) p(t) + \frac{\partial \sigma}{\partial x}(t, t) q(t) + \lambda(t) \frac{\partial g}{\partial x}(t, t) \\
&+ \int_{\mathbb{R}} \frac{\partial \theta}{\partial x}(t, t, e) r(t, e) \nu(de) + \int_0^t \frac{\partial^2 g}{\partial s \partial x}(s, t) \lambda(s) ds \\
&+ \int_t^T \frac{\partial^2 b}{\partial s \partial x}(s, t) p(s) ds + \int_t^T \frac{\partial^2 \sigma}{\partial s \partial x}(s, t) \mathbb{E}[D_t p(s) | \mathcal{F}_t] ds \\
&+ \int_{\mathbb{R}} \frac{\partial^2 \theta}{\partial s \partial x}(s, t, e) \mathbb{E}[D_{t,e} p(s) | \mathcal{F}_t] \nu(de) ds \\
&+ \int_0^t \frac{\partial^2 g}{\partial x \partial z}(s, t) \frac{\partial Z}{\partial s}(s, t) \lambda(s) ds \\
&+ \int_0^t \frac{\partial}{\partial x}(\nabla_k g(s, t)) \left(\frac{\partial K}{\partial s}(s, t, \cdot) \right) ds, \tag{2.39}
\end{aligned}$$

and that

$$\begin{aligned}
\nabla_k \mathcal{H}(t) &= \nabla_k g(t, t) \lambda(t) + \int_0^t \nabla_k \left(\frac{\partial}{\partial s} g(s, t) \right) \lambda(s) ds \\
&+ \int_0^t \nabla_k \left(\frac{\partial g}{\partial z}(s, t) \right) \frac{\partial Z}{\partial s}(s, t) \lambda(s) ds \\
&+ \int_0^t \nabla_k^2 g(s, t) \frac{\partial K}{\partial s}(s, t, \cdot) \lambda(s) ds, \tag{2.40}
\end{aligned}$$

similarly for $\frac{\partial \mathcal{H}}{\partial y}(t)$ and $\frac{\partial \mathcal{H}}{\partial z}(t)$. Combining (2.36) – (2.38) with (2.35), (2.39) – (2.40) and by the definition of μ , we obtain

$$\frac{d}{d\epsilon} J(u + \epsilon \mu) |_{\epsilon=0} = \mathbb{E} \left[\int_0^T \frac{\partial \mathcal{H}}{\partial u}(t) \mu(t) dt \right] = \mathbb{E} \left[\int_t^{t+\epsilon} \frac{\partial \mathcal{H}}{\partial u}(s) ds \alpha \right].$$

We conclude that

$$\frac{d}{d\epsilon} J(u + \epsilon\mu) \big|_{\epsilon=0} = 0$$

if and only if

$$\mathbb{E} \left[\frac{\partial \mathcal{H}}{\partial u}(t) \mid \mathcal{G}_t \right] = 0.$$

□

3 Existence and uniqueness of solutions of BSVIE

In order to prove the existence the uniqueness of the backward stochastic Volterra integral equations (BSVIE), let us introduce the following BSVIE in the unknown Y, Z and K :

$$\begin{aligned} Y(t) = \psi(t) + \int_t^T g(t, s, Y(s), Z(t, s), K(t, s, \cdot)) ds - \int_t^T Z(t, s) dB(s) \\ - \int_t^T \int_{\mathbb{R}} K(t, s, e) \tilde{N}(ds, de), t \in [0, T]. \end{aligned} \quad (3.1)$$

In this section we prove existence and uniqueness of solutions of (3.1), following the approach by Yong [16] and [17], but now we have jumps. The papers by Wang and Zhang [14], and by Ren [9] studied more general cases of (3.1) and our case can be seen as a particular case of theirs, but we have included this part because it will be more convenient for the reader to have a direct and simple approach. For related results on BSVIE, we refer to Shi and Wang [11]-[10].

Let us now introduce the following spaces:

For any $\beta \geq 0$, let $\Delta := \{(t, s) \in [0, T]^2 : t \leq s\}$ and $H_{\Delta}^{2,\beta}[0, T]$ be a space of all processes (Y, Z, K) , such that $Y : [0, T] \times \Omega \rightarrow \mathbb{R}$ is \mathbb{F} -adapted, and $Z : \Delta \times \Omega \rightarrow \mathbb{R}$, $K : \Delta \times \mathbb{R}_0 \times \Omega \rightarrow \mathbb{R}$ with $s \rightarrow Z(t, s)$ and $s \rightarrow K(t, s, \cdot)$ being \mathbb{F} -adapted on $[t, T]$, equipped with the norm

$$\begin{aligned} \|(Y, Z, K)\|_{H_{\Delta}^{2,\beta}[0, T]}^2 \\ := \mathbb{E} \int_0^T \left[e^{\beta t} |Y(t)|^2 + \int_t^T e^{\beta s} |Z(t, s)|^2 ds + \int_t^T \int_{\mathbb{R}} e^{\beta s} |K(t, s, e)|^2 \nu(ds, de) \right] dt. \end{aligned} \quad (3.2)$$

Clearly $H_{\Delta}^{2,\beta}[0, T]$ is a Hilbert space. It is easy to see that for any $\beta > 0$, the norm $\|\cdot\|_{H_{\Delta}^{2,\beta}[0, T]}$ is equivalent to $\|\cdot\|_{H_{\Delta}^{2,0}[0, T]}$ obtained from $\|\cdot\|_{H_{\Delta}^{2,\beta}[0, T]}$ by taking $\beta = 0$. We now make the following assumptions:

Assumptions (H.1)

- The function $g : [0, T]^2 \times \mathbb{R}^3 \times L^2(\nu) \times \Omega \rightarrow \mathbb{R}$, is such that

$$1. \mathbb{E} \left[\int_0^T \left(\int_t^T g(t, s, 0, 0, 0) ds \right)^2 dt \right] < +\infty,$$

2. There exists a constant $c > 0$, such that, for all $t, s \in [0, T]$

$$\begin{aligned} & |g(t, s, y, z, k(\cdot)) - g(t, s, y', z', k'(\cdot))| \\ & \leq c \left(|y - y'| + |z - z'| + \left(\int_{\mathbb{R}} |k(e) - k'(e)|^2 \nu(de) \right)^{\frac{1}{2}} \right) \end{aligned}$$

for all $y, y', z, z', k(\cdot), k'(\cdot)$.

- $\psi(\cdot) \in L^2_{\mathcal{F}_T}(\Omega, \mathbb{R})$.

Theorem 3.1. *Under the assumptions (H.1), there exists a unique solution $(Y, Z, K) \in H^{2,\beta}_{\Delta}[0, T]$ of the BSVIE (3.1).*

Proof For a given triple of processes $(y(\cdot), z(\cdot, \cdot), k(\cdot, \cdot, \cdot)) \in H^{2,\beta}_{\Delta}[0, T]$, consider the following simple BSVIE in the unknown triple (Y, Z, K) :

$$Y(t) = \psi(t) + \int_t^T \bar{g}(t, s) ds - \int_t^T Z(t, s) dB(s) - \int_t^T \int_{\mathbb{R}} K(t, s, e) \tilde{N}(ds, de), \quad (3.3)$$

where we denote by

$$\bar{g}(t, s) = g(t, s, y(s), z(t, s), k(t, s, \cdot)), \text{ for } (t, s) \in \Delta.$$

To solve (3.3) for (Y, Z, K) , we introduce the following family of BSDE (parameterized by $t \in [0, T]$):

$$\chi(r, t) = \psi(t) + \int_r^T \bar{g}(t, s) ds - \int_r^T \eta(s, t) dB(s) - \int_r^T \int_{\mathbb{R}} \xi(s, t, e) \tilde{N}(ds, de), \quad r \in (t, T],$$

It is well-known that the above BSDE admits a unique adapted solution $(\chi(\cdot, t), \eta(\cdot, t), \xi(\cdot, t, \cdot))$ and the following estimate holds:

$$\begin{aligned} & \mathbb{E} \left[\sup_{r \in [t, T]} |\chi(r, t)|^2 + \int_t^T |\eta(s, t)|^2 ds + \int_t^T \int_{\mathbb{R}} |\xi(s, t, e)|^2 \nu(de) ds \right] \\ & \leq C \mathbb{E} \left[|\psi(t)|^2 + \left(\int_t^T \bar{g}(t, s) ds \right)^2 \right]. \end{aligned}$$

Now let

$$Y(t) = \chi(t, t), \quad Z(t, s) = \eta(s, t), \quad K(t, s, \cdot) = \xi(s, t, \cdot), \text{ for all } (t, s) \in \Delta.$$

Then $(Y(\cdot), Z(\cdot, \cdot), K(\cdot, \cdot, \cdot))$ is an adapted solution to the BSVIE (3.3), and

$$\begin{aligned} & \mathbb{E} \left[|Y(t)|^2 + \int_t^T |Z(t, s)|^2 ds + \int_t^T \int_{\mathbb{R}} |K(t, s, e)|^2 \nu(de) ds \right] \\ & = \mathbb{E} \left[\left| \psi(t) + \int_t^T \bar{g}(t, s) ds \right|^2 \right] \\ & \leq 2 \mathbb{E} \left[|\psi(t)|^2 + \left(\int_t^T \bar{g}(t, s) ds \right)^2 \right]. \end{aligned}$$

Therefore, by integrating both sides of the inequality above, we get

$$\begin{aligned} & \mathbb{E} \left[\int_0^T \left(|Y(t)|^2 + \int_t^T |Z(t,s)|^2 ds + \int_t^T \int_{\mathbb{R}} |K(t,s,e)|^2 \nu(de) ds \right) dt \right] \\ & \leq 2\mathbb{E} \int_0^T \left[|\psi(t)|^2 + \left(\int_t^T \bar{g}(t,s) ds \right)^2 \right] dt. \end{aligned}$$

Adding and subtracting $g(t,s,0,0,0)$ on the left side, then by the Lipschitz assumption, we obtain

$$\begin{aligned} & \mathbb{E} \left[\int_0^T \left(|Y(t)|^2 + \int_t^T |Z(t,s)|^2 ds + \int_t^T \int_{\mathbb{R}} |K(t,s,e)|^2 \nu(de) ds \right) dt \right] \\ & \leq C\mathbb{E} \int_0^T \left[|\psi(t)|^2 + \left(\int_t^T g(t,s,0,0,0) ds \right)^2 \right] dt \\ & + C\mathbb{E} \left[\int_0^T \left(|y(t)|^2 + \int_t^T |z(s)|^2 ds + \int_t^T \int_{\mathbb{R}} |k(t,s,e)|^2 \nu(de) ds \right) dt \right], \end{aligned}$$

for some constant C . Thus, $(y,z,k) \mapsto (Y,Z,K)$ defines a map from $H_{\Delta}^{2,\beta}[0,T]$ to itself.

Now, we want to prove that this mapping is contracting in $H_{\Delta}^{2,\beta}[0,T]$ under the norm $\|\cdot\|_{H_{\Delta}^{2,\beta}[0,T]}$. We show that if for $i = 1, 2$, $(y_i, z_i, k_i) \in H_{\Delta}^{2,\beta}[0,T]$ and (Y_i, Z_i, K_i) is the corresponding adapted solution to equation (3.1), then

$$\begin{aligned} & \mathbb{E} \left[\int_0^T \left(e^{\beta t} |Y_1(t) - Y_2(t)|^2 + \int_t^T e^{\beta s} |Z_1(t,s) - Z_2(t,s)|^2 ds \right. \right. \\ & \quad \left. \left. + \int_t^T e^{\beta s} \int_{\mathbb{R}} |K_1(t,s,e) - K_2(t,s,e)|^2 \nu(de) ds \right) dt \right] \\ & \leq \frac{C}{\beta} \mathbb{E} \left[\int_0^T \left(e^{\beta t} |y_1(t) - y_2(t)|^2 + \int_t^T e^{\beta s} |z_1(t,s) - z_2(t,s)|^2 ds \right. \right. \\ & \quad \left. \left. + \int_t^T e^{\beta s} \int_{\mathbb{R}} |k_1(t,s,e) - k_2(t,s,e)|^2 \nu(de) ds \right) dt \right], \end{aligned}$$

which means that

$$\|(Y, Z, K)\|_{H_{\Delta}^{2,\beta}[0,T]}^2 \leq \frac{C}{\beta} \|(y, z, k)\|_{H_{\Delta}^{2,\beta}[0,T]}^2.$$

Hence, the mapping $(y, z, k) \mapsto (Y, Z, K)$ is contracting on $H_{\Delta}^{2,\beta}[0,T]$ for large enough $\beta > 0$. Then, (Y, Z, K) is a unique solution for the BSVIE (3.1). \square

4 Application: Optimal recursive utility consumption

As an illustration of our general results above, we now apply them to solve the optimal recursive utility consumption problem (1.5) described in the Introduction. Our example

is related to the examples discussed in [3] and [8], but now the cash flow is modelled by a stochastic Volterra equation and the utility is represented by the recursive utility. As pointed out after (1.2) in the Introduction, the Volterra equation contains history terms and can therefore be viewed as a model for a system with memory. Thus, we assume that the cash flow $X(t) = X^c(t)$ being exposed to a \mathbb{G} -adapted consumption rate $c(t)$, satisfies the stochastic Volterra equation

$$X(t) = \xi + \int_0^t (\alpha(t, s) - c(s)) X(s) ds + \int_0^t \beta(t, s) X(s) dB(s) + \int_0^t \int_{\mathbb{R}} \pi(t, s, e) X(s) \tilde{N}(ds, de), t \in [0, T], \quad (4.1)$$

where we assume for simplicity that ξ is a (deterministic) constant and $\alpha, \beta : [0, T]^2 \rightarrow \mathbb{R}$ and $\pi : [0, T]^2 \times \mathbb{R}_0 \rightarrow \mathbb{R}$ are deterministic functions with α, β and π bounded. The FSVIE (4.1) can be written in its differential form as

$$\begin{aligned} dX(t) &= (\alpha(t, t) - c(t)) X(t) dt + \left(\int_0^t \frac{\partial \alpha}{\partial t}(t, s) X(s) ds \right) dt \\ &+ \beta(t, t) X(t) dB(t) + \left(\int_0^t \frac{\partial \beta}{\partial t}(t, s) X(s) dB(s) \right) dt \\ &+ \int_{\mathbb{R}} \pi(t, t, e) X(t) \tilde{N}(dt, de) + \left(\int_{\mathbb{R}} \int_0^t \frac{\partial \pi}{\partial t}(t, s, e) X(s) \tilde{N}(ds, de) \right) dt, t \in [0, T]. \end{aligned}$$

The recursive utility process $Y(t)$ of Duffie and Epstein [5] has the following linear form

$$\begin{aligned} dY(t) &= -[\gamma(t)Y(t) + \ln c(t)X(t)] dt + Z(t)dB(t) \\ &+ \int_{\mathbb{R}} K(t, e) \tilde{N}(dt, de), t \in [0, T]. \end{aligned} \quad (4.2)$$

Our problem (1.5) is to maximise the performance functional

$$J(c) := Y^c(0) \quad (4.3)$$

over all control processes $c \in \mathcal{U}_{\mathbb{G}}$, where in this case $\mathcal{U}_{\mathbb{G}}$ is the set of all \mathbb{G} -adapted nonnegative processes.

This problem is a special case of the problem discussed in the previous sections, with $f = 0$, $\varphi = 0$ and $\psi(y) = y$. The Hamiltonian associated to our problem is defined by

$$\begin{aligned} \mathcal{H}(t, s, x, y, p, q) &= (\alpha(t, t) - c(t)) px + \int_t^T \frac{\partial \alpha}{\partial s}(s, t) x(s) p(s) ds \\ &+ \beta(t, t) qx + \int_t^T \frac{\partial \beta}{\partial s}(s, t) x(s) \mathbb{E}[D_t p(s) | \mathcal{F}_t] ds \\ &+ \int_{\mathbb{R}} \pi(t, t, e) x r(t, e) \nu(de) \\ &+ \int_{\mathbb{R}} \int_t^T \frac{\partial \pi}{\partial s}(s, t, e) x(s) \mathbb{E}[D_{t,e} p(s) | \mathcal{F}_t] \nu(de) ds \\ &+ [\gamma(t)y + \ln c(t) + \ln x] \lambda. \end{aligned}$$

The corresponding backward-forward system for the adjoint processes (p, q, r) and λ are

$$\begin{cases} dp(t) = - \left[(\alpha(t, t) - c(t)) p(t) + \int_t^T \frac{\partial \alpha}{\partial s}(s, t) p(s) ds \right. \\ \quad + \beta(t, t) q(t) + \int_t^T \frac{\partial \beta}{\partial s}(s, t) \mathbb{E} [D_t p(s) | \mathcal{F}_t] ds + \int_{\mathbb{R}} \pi(t, t, e) r(t, e) \nu(de) \\ \quad \left. + \int_{\mathbb{R}} \int_t^T \frac{\partial \pi}{\partial s}(s, t, e) \mathbb{E} [D_{t,e} p(s) | \mathcal{F}_t] \nu(ds, de) + \frac{\lambda(t)}{X(t)} \right] dt \\ \quad + q(t) dB(t) + \int_{\mathbb{R}} r(t, e) \tilde{N}(dt, de), t \in [0, T], \\ p(T) = 0, \end{cases} \quad (4.4)$$

and

$$d\lambda(t) = \gamma(t)\lambda(t)dt, t \in [0, T], \lambda(0) = 1. \quad (4.5)$$

The solution of the differential equation (4.5) is

$$\lambda(t) = \exp \left(- \int_0^t \gamma(s) ds \right), t \in [0, T].$$

Now, maximising the Hamiltonian w.r.t c gives the first order condition

$$c(t) = \mathbb{E} \left[\frac{\lambda(t)}{p(t)X(t)} \middle| \mathcal{G}_t \right] \quad t \in [0, T]. \quad (4.6)$$

Applying Itô's formula, we get

$$\begin{aligned} d(p(t)X(t)) &= p(t)dX(t) + X(t)dp(t) + d[p(t)X(t)] \\ &= p(t) \left\{ (\alpha(t, t) - c(t)) X(t)dt + \left(\int_0^t \frac{\partial \alpha}{\partial t}(t, s) X(s) ds \right) dt \right. \\ &\quad + \beta(t, t) X(t) dB(t) + \left(\int_0^t \frac{\partial \beta}{\partial t}(t, s) X(s) dB(s) \right) dt \\ &\quad \left. + \int_{\mathbb{R}} \pi(t, t, e) X(t) \tilde{N}(dt, de) + \left(\int_{\mathbb{R}} \int_0^t \frac{\partial \pi}{\partial t}(t, s, e) X(s) \tilde{N}(ds, de) \right) dt \right\} \\ &\quad - X(t) \left\{ (\alpha(t, t) - c(t)) p(t)dt + \left(\int_t^T \frac{\partial \alpha}{\partial s}(s, t) p(s) ds \right) dt + \beta(t, t) q(t)dt \right. \\ &\quad + \left(\int_t^T \frac{\partial \beta}{\partial s}(s, t) \mathbb{E} [D_t p(s) | \mathcal{F}_t] ds \right) dt + \int_{\mathbb{R}} \pi(t, t, e) r(t, e) \nu(dt, de) \\ &\quad + \left(\int_{\mathbb{R}} \int_t^T \frac{\partial \pi}{\partial s}(s, t, e) \mathbb{E} [D_{t,e} p(s) | \mathcal{F}_t] \nu(de) ds \right) dt + \frac{\lambda(t)}{X(t)} dt \\ &\quad \left. + q(t) dB(t) + \int_{\mathbb{R}} r(t, e) \tilde{N}(dt, de) \right\} \\ &\quad + \beta(t, t) X(t) q(t) dt + \int_{\mathbb{R}} \pi(t, t, e) X(t) r(t, e) \nu(dt, de). \end{aligned}$$

Collecting the terms, we see that the above reduces to

$$\begin{cases} p(t)X(t) = p(0)X(0) - \int_0^t \lambda(s) ds \\ \quad + \int_0^t \{p(s)X(s)\beta(s, s) - X(s)q(s)\} dB(s) \\ \quad + \int_0^t \{p(s)X(s)\pi(s, s, e) - X(s)r(s, e)\} \tilde{N}(ds, de), t \in [0, T], \\ p(T)X(T) = 0. \end{cases}$$

Therefore, if we define

$$\begin{aligned} u(t) &= p(t)X(t) \\ v(t) &= p(s)X(s)\beta(s, s) - X(s)q(s) \\ w(t, e) &= p(s)X(s)\pi(s, s, e) - X(s)r(s, e), \end{aligned}$$

then (u, v, w) solves the linear BSDE

$$\begin{cases} du(t) = -\lambda(t)dt + v(t)dB(t) + \int_{\mathbb{R}} w(t, e)\tilde{N}(dt, de), t \in [0, T], \\ u(T) = 0. \end{cases}$$

The solution of this linear BSDE is

$$u(t) = \mathbb{E} \left[\int_t^T \lambda(s)ds \middle| \mathcal{F}_t \right] = p(t)X(t). \quad (4.7)$$

Combined with (4.6) this gives

$$c(t) = c^*(t) = \mathbb{E} \left[\frac{\exp \left(- \int_0^t \gamma(s)ds \right)}{\mathbb{E} \left[\int_t^T \exp \left(- \int_0^s \gamma(r)dr \right) ds \middle| \mathcal{F}_t \right]} \middle| \mathcal{G}_t \right]. \quad (4.8)$$

In particular, since $\lambda > 0$ by (4.5) we get that $p(t)X(t) > 0$. Thus we see that $c(t)$ is well-defined in (4.6) and $c^*(t) > 0$ for all $t \in [0, T]$. Therefore $c^* \in \mathcal{U}_{\mathbb{G}}$, and we conclude that c^* is indeed optimal. We have proved

Theorem 4.1. *The optimal recursive utility consumption rate $c^*(t)$ for the problem (1.5) (with ξ constant) is given by (4.8).*

5 Appendix

5.1 Some basic concepts from Banach space theory

To explain the notation used in this paper, we briefly recall some basic concepts from Banach space theory:

Let \mathcal{X}, \mathcal{Y} be two Banach spaces with norms $\|\cdot\|_{\mathcal{X}}, \|\cdot\|_{\mathcal{Y}}$, respectively, and let $F : \mathcal{X} \rightarrow \mathcal{Y}$.

- We say that F has a directional derivative (or Gâteaux derivative) at $v \in \mathcal{X}$ in the direction $w \in \mathcal{X}$ if

$$D_w F(v) := \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} (F(v + \varepsilon w) - F(v))$$

exists.

- We say that F is Fréchet differentiable at $v \in \mathcal{X}$ if there exists a continuous linear map $A : \mathcal{X} \rightarrow \mathcal{Y}$ such that

$$\lim_{\substack{h \rightarrow 0 \\ h \in \mathcal{X}}} \frac{1}{\|h\|_{\mathcal{X}}} \|F(v + h) - F(v) - A(h)\|_{\mathcal{Y}} = 0.$$

In this case we call A the *gradient* (or Fréchet derivative) of F at v and we write

$$A = \nabla_v F.$$

- If F is Fréchet differentiable at v with Fréchet derivative $\nabla_v F$, then F has a directional derivative in all directions $w \in \mathcal{X}$ and

$$D_w F(v) := \langle \nabla_v F, w \rangle = \nabla_v F(w) = \nabla_v F w.$$

In particular, note that if F is a linear operator, then $\nabla_v F = F$ for all v .

5.2 A brief review of Hida-Malliavin calculus for Lévy processes

For the convenience of the reader, in this section we recall the basic definition and properties of Hida-Malliavin calculus for Lévy processes related to this paper. The following summary is based on [2]. A general reference for this presentation is the book [4].

First, recall the Lévy-Itô decomposition theorem, which states that any Lévy process $Y(t)$ with

$$\mathbb{E}[Y^2(t)] < \infty \quad \text{for all } t$$

can be written

$$Y(t) = at + bB(t) + \int_0^t \int_{\mathbb{R}} \zeta \tilde{N}(ds, d\zeta)$$

with constants a and b . In view of this we see that it suffices to deal with Hida-Malliavin calculus for $B(\cdot)$ and for

$$\eta(\cdot) := \int_0^\cdot \int_{\mathbb{R}} \zeta \tilde{N}(ds, d\zeta)$$

separately.

5.3 Hida-Malliavin calculus for $B(\cdot)$

A natural starting point is the Wiener-Itô chaos expansion theorem, which states that any $F \in L^2(\mathcal{F}_T, P)$ can be written

$$F = \sum_{n=0}^{\infty} I_n(f_n) \tag{5.1}$$

for a unique sequence of symmetric deterministic functions $f_n \in L^2(\lambda^n)$, where λ is Lebesgue measure on $[0, T]$ and

$$I_n(f_n) = n! \int_0^T \int_0^{t_n} \cdots \int_0^{t_2} f_n(t_1, \dots, t_n) dB(t_1) dB(t_2) \cdots dB(t_n) \tag{5.2}$$

(the n -times iterated integral of f_n with respect to $B(\cdot)$) for $n = 1, 2, \dots$ and $I_0(f_0) = f_0$ when f_0 is a constant.

Moreover, we have the isometry

$$\mathbb{E}[F^2] = \|F\|_{L^2(P)}^2 = \sum_{n=0}^{\infty} n! \|f_n\|_{L^2(\lambda^n)}^2. \tag{5.3}$$

Definition 5.1 (Hida-Malliavin derivative D_t with respect to $B(\cdot)$).

Let $\mathbb{D}_{1,2}^{(B)}$ be the space of all $F \in L^2(\mathcal{F}_T, P)$ such that its chaos expansion (4) satisfies

$$\|F\|_{\mathbb{D}_{1,2}^{(B)}}^2 := \sum_{n=1}^{\infty} nn! \|f_n\|_{L^2(\lambda^n)}^2 < \infty. \quad (5.4)$$

For $F \in \mathbb{D}_{1,2}^{(B)}$ and $t \in [0, T]$, we define the *Hida-Malliavin derivative* or *the stochastic gradient* of F at t (with respect to $B(\cdot)$), $D_t F$, by

$$D_t F = \sum_{n=1}^{\infty} n I_{n-1}(f_n(\cdot, t)), \quad (5.5)$$

where the notation $I_{n-1}(f_n(\cdot, t))$ means that we apply the $(n-1)$ -times iterated integral to the first $n-1$ variables t_1, \dots, t_{n-1} of $f_n(t_1, t_2, \dots, t_n)$ and keep the last variable $t_n = t$ as a parameter.

One can easily check that

$$\mathbb{E} \left[\int_0^T (D_t F)^2 dt \right] = \sum_{n=1}^{\infty} nn! \|f_n\|_{L^2(\lambda^n)}^2 = \|F\|_{\mathbb{D}_{1,2}^{(B)}}^2, \quad (5.6)$$

so $(t, \omega) \rightarrow D_t F(\omega)$ belongs to $L^2(\lambda \times P)$.

Example 5.2. If $F = \int_0^T f(t) dB(t)$ with $f \in L^2(\lambda)$ deterministic, then

$$D_t F = f(t) \text{ for a.a. } t \in [0, T].$$

More generally, if $u(s)$ is Skorohod integrable, $u(s) \in \mathbb{D}_{1,2}$ for a.a. s and $D_t u(s)$ is Skorohod integrable for a.a. t , then

$$D_t \left(\int_0^T u(s) \delta B(s) \right) = \int_0^T D_t u(s) \delta B(s) + u(t) \text{ for a.a. } (t, \omega), \quad (5.7)$$

where $\int_0^T \psi(s) \delta B(s)$ denotes the Skorohod integral of a process ψ with respect to $B(\cdot)$.

Some other basic properties of the Hida-Malliavin derivative D_t are the following:

(i) **Chain rule**

Suppose $F_1, \dots, F_m \in \mathbb{D}_{1,2}^{(B)}$ and that $\psi : \mathbb{R}^m \rightarrow \mathbb{R}$ is C^1 with bounded partial derivatives. Then, $\psi(F_1, \dots, F_m) \in \mathbb{D}_{1,2}$ and

$$D_t \psi(F_1, \dots, F_m) = \sum_{i=1}^m \frac{\partial \psi}{\partial x_i}(F_1, \dots, F_m) D_t F_i. \quad (5.8)$$

(ii) **Duality formula**

Suppose $u(t)$ is \mathcal{F}_t -adapted with $\mathbb{E}[\int_0^T u^2(t)dt] < \infty$ and let $F \in \mathbb{D}_{1,2}^{(B)}$. Then,

$$\mathbb{E}[F \int_0^T u(t)dB(t)] = \mathbb{E}[\int_0^T u(t)D_t F dt]. \quad (5.9)$$

(iii) **Malliavin derivative and adapted processes**

If φ is an \mathbb{F} -adapted process, then

$$D_s \varphi(t) = 0 \text{ for } s > t.$$

Remark 5.3. We put $D_t \varphi(t) = \lim_{s \rightarrow t-} D_s \varphi(t)$ (if the limit exists).

Remark 5.4. It was proved in [1] that one can extend the Hida-Malliavin derivative operator D_t from $\mathbb{D}_{1,2}$ to all of $L^2(\mathcal{F}_T, P)$ in such a way that, also denoting the extended operator by D_t , for all $F \in L^2(\mathcal{F}_T, P)$ we have

$$D_t F \in (\mathcal{S})^* \text{ and } (t, \omega) \mapsto \mathbb{E}[D_t F \mid \mathcal{F}_t] \text{ belongs to } L^2(\lambda \times P) \quad (5.10)$$

Here $(\mathcal{S})^*$ is the Hida space of stochastic distributions.

Moreover, the following generalized Clark-Haussmann-Ocone formula was proved:

$$F = \mathbb{E}[F] + \int_0^T \mathbb{E}[D_t F \mid \mathcal{F}_t] dB(t) \quad (5.11)$$

for all $F \in L^2(\mathcal{F}_T, P)$. See Theorem 3.11 in [1] and also Theorem 6.35 in [4].

We can use this to get the following extension of the duality formula (5.9):

Proposition 5.5. The generalized duality formula

Let $F \in L^2(\mathcal{F}_T, P)$ and let $\varphi(t, \omega) \in L^2(\lambda \times P)$ be adapted. Then

$$\mathbb{E}[F \int_0^T \varphi(t)dB(t)] = \mathbb{E}[\int_0^T \mathbb{E}[D_t F \mid \mathcal{F}_t] \varphi(t)dt]. \quad (5.12)$$

Proof By (5.10) and (5.11) and the Itô isometry we get

$$\begin{aligned} \mathbb{E}[F \int_0^T \varphi(t)dB(t)] &= \mathbb{E}[(\mathbb{E}[F] + \int_0^T \mathbb{E}[D_t F \mid \mathcal{F}_t] dB(t))(\int_0^T \varphi(t)dB(t))] \\ &= \mathbb{E}[\int_0^T \mathbb{E}[D_t F \mid \mathcal{F}_t] \varphi(t)dt]. \end{aligned} \quad (5.13)$$

□

We will use this extension of the Hida-Malliavin derivative from now on.

5.4 Hida-Malliavin calculus for $\tilde{N}(\cdot)$

The construction of a stochastic derivative/Hida-Malliavin derivative in the pure jump martingale case follows the same lines as in the Brownian motion case. In this case, the corresponding Wiener-Itô Chaos Expansion Theorem states that any $F \in L^2(\mathcal{F}_T, P)$ (where, in this case, $\mathcal{F}_t = \mathcal{F}_t^{(\tilde{N})}$ is the σ -algebra generated by $\eta(s) := \int_0^s \int_{\mathbb{R}} \zeta \tilde{N}(dr, d\zeta)$; $0 \leq s \leq t$) can be written as

$$F = \sum_{n=0}^{\infty} I_n(f_n); \quad f_n \in \hat{L}^2((\lambda \times \nu)^n), \quad (5.14)$$

where $\hat{L}^2((\lambda \times \nu)^n)$ is the space of functions $f_n(t_1, \zeta_1, \dots, t_n, \zeta_n)$; $t_i \in [0, T]$, $\zeta_i \in \mathbb{R}_0$ such that $f_n \in L^2((\lambda \times \nu)^n)$ and f_n is symmetric with respect to the pairs of variables $(t_1, \zeta_1), \dots, (t_n, \zeta_n)$.

It is important to note that in this case, the n -times iterated integral $I_n(f_n)$ is taken with respect to $\tilde{N}(dt, d\zeta)$ and not with respect to $d\eta(t)$. Thus, we define

$$I_n(f_n) := n! \int_0^T \int_{\mathbb{R}_0} \int_0^{t_n} \int_{\mathbb{R}_0} \cdots \int_0^{t_2} \int_{\mathbb{R}_0} f_n(t_1, \zeta_1, \dots, t_n, \zeta_n) \tilde{N}(dt_1, d\zeta_1) \cdots \tilde{N}(dt_n, d\zeta_n) \quad (5.15)$$

for $f_n \in \hat{L}^2((\lambda \times \nu)^n)$.

The Itô isometry for stochastic integrals with respect to $\tilde{N}(dt, d\zeta)$ then gives the following isometry for the chaos expansion:

$$\|F\|_{L^2(P)}^2 = \sum_{n=0}^{\infty} n! \|f_n\|_{L^2((\lambda \times \nu)^n)}^2. \quad (5.16)$$

As in the Brownian motion case, we use the chaos expansion to define the Malliavin derivative. Note that in this case, there are two parameters t, ζ , where t represents time and $\zeta \neq 0$ represents a generic jump size.

Definition 5.6 (Hida-Malliavin derivative $D_{t,\zeta}$ with respect to $\tilde{N}(\cdot, \cdot)$). [4] Let $\mathbb{D}_{1,2}^{(\tilde{N})}$ be the space of all $F \in L^2(\mathcal{F}_T, P)$ such that its chaos expansion (5.14) satisfies

$$\|F\|_{\mathbb{D}_{1,2}^{(\tilde{N})}}^2 := \sum_{n=1}^{\infty} n n! \|f_n\|_{L^2((\lambda \times \nu)^2)}^2 < \infty. \quad (5.17)$$

For $F \in \mathbb{D}_{1,2}^{(\tilde{N})}$, we define the Hida-Malliavin derivative of F at (t, ζ) (with respect to $\tilde{N}(\cdot)$), $D_{t,\zeta}F$, by

$$D_{t,\zeta}F := \sum_{n=1}^{\infty} n I_{n-1}(f_n(\cdot, t, \zeta)), \quad (5.18)$$

where $I_{n-1}(f_n(\cdot, t, \zeta))$ means that we perform the $(n-1)$ -times iterated integral with respect to \tilde{N} to the first $n-1$ variable pairs $(t_1, \zeta_1), \dots, (t_n, \zeta_n)$, keeping $(t_n, \zeta_n) = (t, \zeta)$ as a parameter.

In this case, we get the isometry.

$$\mathbb{E}\left[\int_0^T \int_{\mathbb{R}_0} (D_{t,\zeta} F)^2 \nu(d\zeta) dt\right] = \sum_{n=0}^{\infty} n n! \|f_n\|_{L^2((\lambda \times \nu)^n)}^2 = \|F\|_{\mathbb{D}_{1,2}^{(\tilde{N})}}^2. \quad (5.19)$$

(Compare with (5.6)).

Example 5.7. If $F = \int_0^T \int_{\mathbb{R}} f(t, \zeta) \tilde{N}(dt, d\zeta)$ for some deterministic $f(t, \zeta) \in L^2(\lambda \times \nu)$, then

$$D_{t,\zeta} F = f(t, \zeta) \text{ for } a.a. (t, \zeta).$$

More generally, if $\psi(s, \zeta)$ is Skorohod integrable with respect to $\tilde{N}(\delta s, d\zeta)$, $\psi(s, \zeta) \in \mathbb{D}_{1,2}^{(\tilde{N})}$ for $a.a. s, \zeta$ and $D_{t,z}\psi(s, \zeta)$ is Skorohod integrable for $a.a. (t, z)$, then

$$D_{t,z}\left(\int_0^T \int_{\mathbb{R}} \psi(s, \zeta) \tilde{N}(\delta s, d\zeta)\right) = \int_0^T \int_{\mathbb{R}} D_{t,z}\psi(s, \zeta) \tilde{N}(\delta s, d\zeta) + u(t, z) \text{ for } a.a. t, z, \quad (5.20)$$

where $\int_0^T \int_{\mathbb{R}} \psi(s, \zeta) \tilde{N}(\delta s, d\zeta)$ denotes the *Skorohod integral* of ψ with respect to $\tilde{N}(\cdot, \cdot)$. (See [4] for a definition of such Skorohod integrals and for more details.)

The properties of $D_{t,\zeta}$ corresponding to those of D_t are the following:

(i) **Chain rule** [4]

Suppose $F_1, \dots, F_m \in \mathbb{D}_{1,2}^{(\tilde{N})}$ and that $\phi : \mathbb{R}^m \rightarrow \mathbb{R}$ is continuous and bounded. Then, $\phi(F_1, \dots, F_m) \in \mathbb{D}_{1,2}^{(\tilde{N})}$ and

$$D_{t,\zeta}\phi(F_1, \dots, F_m) = \phi(F_1 + D_{t,\zeta}F_1, \dots, F_m + D_{t,\zeta}F_m) - \phi(F_1, \dots, F_m). \quad (5.21)$$

(ii) **Duality formula** [4]

Suppose $\Psi(t, \zeta)$ is \mathcal{F}_t -adapted and $\mathbb{E}[\int_0^T \int_{\mathbb{R}} \psi^2(t, \zeta) \nu(d\zeta) dt] < \infty$ and let $F \in \mathbb{D}_{1,2}^{(\tilde{N})}$. Then,

$$\mathbb{E}\left[F \int_0^T \int_{\mathbb{R}_0} \Psi(t, \zeta) \tilde{N}(dt, d\zeta)\right] = \mathbb{E}\left[\int_0^T \int_{\mathbb{R}} \Psi(t, \zeta) D_{t,\zeta} F \nu(d\zeta) dt\right]. \quad (5.22)$$

(iii) **Hida-Malliavin derivative and adapted processes** [4]

If φ is an \mathbb{F} -adapted process, then,

$$D_{s,\zeta}\varphi(t) = 0 \text{ for all } s > t.$$

Remark 5.8. We put $D_{t,\zeta}\varphi(t) = \lim_{s \rightarrow t-} D_{s,\zeta}\varphi(t)$ (if the limit exists).

Remark 5.9. As in Remark 3.2 we note that there is an extension of the Hida-Malliavin derivative $D_{t,\zeta}$ from $\mathbb{D}_{1,2}^{(\tilde{N})}$ to $L^2(\mathcal{F}_t \times P)$ such that the following extension of the duality theorem holds:

Proposition 5.10. Generalized duality formula

Suppose $\Psi(t, \zeta)$ is \mathcal{F}_t -adapted and $\mathbb{E}[\int_0^T \int_{\mathbb{R}} \psi^2(t, \zeta) \nu(d\zeta) dt] < \infty$ and let $F \in L^2(\mathcal{F}_T \times P)$. Then,

$$\mathbb{E}\left[F \int_0^T \int_{\mathbb{R}} \Psi(t, \zeta) \tilde{N}(dt, d\zeta)\right] = \mathbb{E}\left[\int_0^T \int_{\mathbb{R}} \Psi(t, \zeta) \mathbb{E}[D_{t,\zeta} F \mid \mathcal{F}_t] \nu(d\zeta) dt\right]. \quad (5.23)$$

We refer to Theorem 13.26 in [4].

Accordingly, note that from now on we are working with this generalized version of the Hida-Malliavin derivative. We emphasize that this generalized Hida-Malliavin derivative DX exists for all $X \in L^2(P)$ as an element of the Hida stochastic distribution space $(\mathcal{S})^*$, and it has the property that the conditional expectation $\mathbb{E}[DX \mid \mathcal{F}_t]$ belongs to $L^2(\lambda \times P)$, where λ is Lebesgue measure on $[0, T]$. Therefore, when using this generalized Hida-Malliavin derivative, combined with conditional expectation, no assumptions on Hida-Malliavin differentiability in the classical sense are needed; we can work on the whole space of random variables in $L^2(P)$.

6 Acknowledgments

We want to thank Yaozhong Hu and Yanqing Wang helpful comments. We are also grateful to an anonymous referee for a very valuable and comprehensive report, which helped us to improve the paper considerably.

References

- [1] Aase, K., Øksendal, B., Privault, N. and Ubøe, J.: White noise generalizations of the Clark-Haussmann-Ocone theorem with application to mathematical finance. *Finance Stochast.* 4 (2000), 465-496.
- [2] Agram, N., Øksendal, B.: Malliavin calculus and optimal control of stochastic Volterra equations. *J. Optim. Theory Appl.* , DOI 10.1007/s10957-015-0735-5 (2015).
- [3] Agram, N., Øksendal, B.: Infinite horizon optimal control of forward-backward stochastic differential equations with delay. *J. Comput. Appl. Math.* 259, 336–349 (2014)
- [4] Di Nunno, G., Øksendal, B. and Proske, F.: *Malliavin Calculus for Lévy Processes and Applications to Finance*. Corrected, Second Printing, Springer (2009).
- [5] Duffie, D., Epstein, L.G.: Stochastic differential utility, *Econometrica* 60 (2) 353–394 (1992).
- [6] Hu, Y. and Øksendal, B.: Linear backward stochastic Volterra equations. Manuscript August (2016).

- [7] Lin, J.: Adapted solution of a backward stochastic nonlinear Volterra integral equation, *Stochastic Analysis and Applications*, vol. 20, no. 1, pp. 165–183 (2002).
- [ØS1] B. Øksendal and A. Sulem: Risk minimization in financial markets modeled by Itô-Lévy processes. *Afrika Matematika* (2014), DOI: 10.1007/s13370-014-02489-9.
- [8] Øksendal, B., Sulem, A.: Optimal control of predictive mean-field equations and applications to finance. In F.E. Benth and G. Di Nunno (eds.), *Stochastics of Environmental and Financial Economics*, Springer Proceedings in Mathematics and Statistics 138, DOI 10.1007/978-3-319-23425-0_12 (2016).
- [9] Ren, Y.: On solutions of backward stochastic Volterra integral equations with jumps in Hilbert spaces. *J Optim Theory Appl* 144: 319–333 (2010).
- [10] Shi, Y., Wang, T. and Yong, J.: Optimal control problems of forward-backward stochastic Volterra integral equations. arXiv: 1404.7577v1 (2014).
- [11] Shi, Y. and Wang, T.: Solvability of general backward stochastic Volterra integral equations. *J. Korean Math. Soc.* 49 No. 6, pp. 1301–1321 (2012).
- [12] Wang, T. and Shi, Y.: A maximum principle for forward-backward stochastic Volterra integral equations and applications in finance. arXiv 1004.2206v1 (2010).
- [13] Wang, Z. and Zhang, X.: Optimal control problems of forward-backward stochastic Volterra integral equations with closed control regions. arXiv:1602.05661v1 (2016).
- [14] Wang, Z. and Zhang, X.: Non-Lipschitz backward stochastic Volterra type equations with jumps. *Stoch. Dyn.* 07:479-496 (2007).
- [15] Wei, Q. and Xiao, X.: An optimal control problem of forward-backward stochastic Volterra integral equations with state constraints. *Abstract and Applied Analysis*, Volume (2014), Article ID 432718, 16 (2014).
- [16] Yong, Y.: Backward stochastic Volterra equations and some related problems, *Stochastic Processes and their Applications*, 116, 779-795 (2006).
- [17] Yong, J.: Backward stochastic Volterra integral equations- a brief survey. *Appl. Math. J. Chinese Univ.* 28(4): 383-394 (2013).